2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress

Evaluation of Transfer Learning for Human Activity Recognition among Different Datasets

Md Shafiqul ISLAM, Tsuyoshi OKITA, Sozo INOUE

Kyushu Institute of Technology, Hibiki 2-4, Wakamatsu-ku, Kitakyushu-shi, Fukuoka, 804-8550, JAPAN
{shafiqul|okita|sozo}@sozolab.jp

Abstract— **Human activity recognition is a potential area of research. For better performance, it requires significant amount of labelled data. Collecting labeled activity data is expensive and time-consuming. To solve this problem, transfer learning has been demonstrated very effective as it gathers knowledge from labeled train data of source domain and transfers that knowledge to target domain, which has little or no labeled data. In this paper, we propose unsupervised transfer learning from source dataset to target dataset, which are completely different in terms of number of users and samples. We have used Maximum Mean Discrepancy (MMD) based transfer learning model and compared with base Convolutional Neural Network (CNN) model. We have used 4 datasets for experiment. We have trained the model on a source dataset and then transferred the model to a target dataset, which has no labels to classify activities. We have found that transfer learning model has achieved better performance compared to the base model.**

Keywords- Activity recognition, transfer learning

I. INTRODUCTION

Human activity recognition (HAR) refers to recognize human actions like sitting, walking, jogging etc. It has a wide area of application in healthcare system, smart home, surveillance systems etc. A major concern of HAR is to improve recognition performance. When the train and test data belong to same dataset, which is performed by same users, then machine learning models yield good result. There are several key challenges in HAR. (1) Collecting labeled data. It is very time consuming. In real life if there is no labeled data or if there is missing data then performance begins to drop. (2) Extracting features. It requires domain expertise to extract features from raw data. (3) Different users perform same actions differently due to age, habits, psychology etc. So same activity data in different datasets varies. (4) Large amount of data. Recently deep learning models show very good performance over traditional machine learning models. Deep learning models requires large amount of data. Therefore, if there is insufficient amount of data then the model might not perform well. (5) Probability distribution mismatch. In real world the probability distribution between source and target domain are not same. Within same dataset, this mismatch is less compare to different source and target domains. Therefore, a model trained on a specific dataset might not work well on a different dataset.

To overcome these challenges, we propose unsupervised transfer learning model for HAR. Several researches have been done on transfer learning for HAR. Some use transfer learning models to transfer knowledge among body parts, some transfer among devices, locations etc. [1–3]. Some researchers used deep learning model with some labeled target domain data to fine tune the model [5].

Here, we are proposing transfer learning model between datasets with unlabeled target domain data. To address the first challenge, (1) labelled data, our proposed transfer learning model works on unlabeled target domain data. The source domain data is labeled and target domain data is unlabeled for our model. Our model is trained on labeled source data and then tested on unlabeled target domain data.

For the second challenge, (2) features extraction, our model is CNN based. So, it will automatically extract features. As a result, manual feature extraction is not required. For the third challenge, (3) doing same actions differently, our model takes source and target data as input and extract features from them. It then measures distances between the feature distribution of source and target data and tries to predict similar type of activities. For the fourth challenge, (4) small amount of data, our model is trained on source domain and then transfers the knowledge to target domain. It performs well even if there is small amount of target domain data. It is also not required to train the model from scratch for the target domain after training on source domain dataset. As a result, training requires less time. Finally (5) probability distribution mismatch. Maximum Mean Discrepancy (MMD) with a deep learning model is used to measure the dissimilarity between source and target domain features.

The main contributions of this paper are as follows:

- We propose widely used MMD based unsupervised transfer learning model, which is trained on labeled source data and tested on unlabeled target domain data
- While using knowledge from source to target dataset sometimes the performance degrades. We have provided explanation in this scenario.

Since our source dataset is labeled so we define the source domain as $D_s = \{x_s,\ y_s\}$. On the other hand, the target dataset is completely different from source dataset. As the target dataset is unlabeled so we define the target domain as $D_t = \{x_t\}$. We want to transfer a deep learning model from D_s to D_t.

We have organized the paper as follows: Section 1 covers the introduction of the paper. In Section 2, we present some related works on activity recognition, transfer learning and background. In Section 3, we present our proposed method. In Section 4, we present the experimental and evaluation of

978-1-7281-3024-8/19/$31.00 ©2019 IEEE
DOI 10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00155

IEEE computer society

results. In Section 5, we present discussion. Finally, we conclude the paper with some future work and conclusion in Section 6.

II. Related Works

In this section, we discuss related work on (i) activity recognition (ii) transfer learning and (iii) transfer learning for HAR.

2.1 Activity Recognition

In pervasive computing [14] human activity recognition is a popular research area. Traditional machine learning [14], [15] and deep learning [16] approaches have been successfully applied to human activity recognition and have achieved good accuracy. Deep learning algorithms have shown better performance than traditional machine learning algorithms as deep learning methods can extract features automatically from the raw sensor data. Recursive Neural Network (RNN) model [17], Long Short Term Memory (LSTM) have been used in human activity recognition and have achieved very good accuracy.

2.2 Transfer Learning

Traditional machine learning model is trained and tested on the data which has same input feature space and same data distribution. If the data distribution between training and test data is changed, the performance of the model drops [18]. Obtaining more train data having similar feature space and data distribution like test data is tough and expensive.

Transfer learning can be used to solve this problem. It means transferring knowledge from source domain to target domain. Transfer learning can be used in many fields like computer vision, natural language processing etc. For example, in image classification, one of the models like VGG16, DenseNet121, ResNet50 etc. can be used on a large image data set and then the model can be transferred to a small image dataset which has smaller similar type of images to make the classifier adaptive [19].

There are different types of transfer learning i.e. instance based, feature based, parameter-based transfer learning etc. Instance based transfer learning is basically a re-weighting technique where the weights of the neurons are recalculated [20]. For parameter-based transfer learning, the model is trained on labeled source domain and then uses clustering in target domain [21]. In feature-based transfer learning, the distances between the features of source domain and target domain is calculated to minimize the dissimilarity and then the feature information is transferred between domains [22].

In this paper, we are using feature-based transfer learning. The features from source and target are extracted and then the distance is calculated. Maximum Mean Discrepancy (MMD) is a widely used measurement technique to measure the differences between source and target features. There are some popular deep transfer learning models which are based on MMD, like Deep Domain Confusion (DDC) [9], Domain Adversarial Neural Network (DANN) [11], Joint Adaptation Network (JAN) [12] etc. These methods use an adaptation layer to reduce the discrepancy between the source and target domain features.

2.3 Transfer Learning for HAR

There have been several researches on transfer learning for human activity recognition. Paper [7] proposed MMD based transfer learning model to transfer from one user to another user. Wang et al. [6] proposed Unsupervised Source Selection for Activity Recognition (USSAR) algorithm to select right source domain and then used Transfer Neural Network for Activity Recognition (TNNAR) to transfer knowledge across different body parts of human. Paper [8] used Heterogeneous Deep Convolutional Neural Network (HDCNN) to transfer knowledge from one device to another. Cross-domain activity recognition using Stratified Transfer Learning method is proposed in paper [10].

In real world, there exists some challenges for human activity recognition using machine-learning methods. In a particular dataset, the users perform different actions almost same way but if the datasets are different then the same actions might be performed in different ways by different users. The sampling rate might also be changed for different datasets. As a result, the distribution of source and target features for a particular activity might hold lots of differences. In the same dataset, this discrepancy is less, compared to different datasets. Therefore, it is challenging for the transfer learning models to minimize this discrepancy which might lead to poor performance.

Another major problem is unlabeled data. In real world, collecting labeled data is time consuming an expensive. Many transfer learning methods have been used with little or no labeled data. It is a challenge to train a model in such a way so that even if dataset is different without labels, the performance does not degrade significantly.

If the datasets are different then proper classification of similar types of tasks might be difficult. For example, suppose there is "walking" activity in two datasets. In the first dataset, the user walks very fast while the user of other dataset walks very slow. When the model is trained on first dataset and then tries to recognize actions of second dataset, then it might not be able to detect the activities properly. When transfer learning model is applied it will get the feature distribution of "walking" activity of second dataset very close to "staying" activity of first dataset. Therefore, the model might classify "walking" activity as "staying" activity which is not right. So when the model is trained and tested on a single dataset, the accuracy becomes high. On the other hand, when the transfer learning model is trained on a dataset and then used to recognize activities of different dataset, the performance might be decreased. Although transfer learning helps to transfer knowledge from source to target domain, its performance might decrease due to these real world challenges.

III. Proposed Method

In this section, we present (i) base model (ii) Maximum Mean Discrepancy (MMD) and (iii) transfer learning model for human activity recognition. The MMD is used in the transfer learning model to measure the dissimilarity between source domain features and target domain features.

3.1 Base Model

We have used 1D Convolutional Neural Network (CNN) as the base model. It consists of 2 convolutional layers, 1 max pooling layer and 2 fully connected layers.

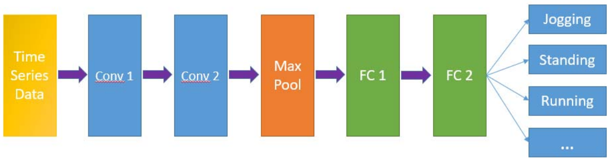

Fig. 1. CNN framework for base model

Fig. 1 shows the overall architecture of the base model that consists of convolutional layers, pooling layers, fully connected layers etc. Each of first and second convolutional layer has 64 filters and kernel size is 1x3. After the second convolutional layer a max pooling layer is used which pool size is 1x2 and stride is 2. In the end 2 fully connected layers FC1 and FC2 are used.

The transfer learning model is developed based on this base model. Therefore, the base model must be very good. The proposed base model is simple in architecture but provides very good performance. This model achieves around 90% accuracy in different datasets like WISDM, HASC, Single Chest etc.

3.2 Maximum Mean Discrepancy (MMD)

Maximum Mean Discrepancy (MMD) means to measure the dissimilarity between 2 probability distributions. It does not require estimating density functions of the distributions initially, which makes MMD an effective criterion. From the kernel Hilbert space, MMD uses functions, which are referred as discriminatory functions.

The two distributions are considered same if the discrepancy is equal to zero. There are several transfer learning algorithms like DAN, DDC which are based on MMD. In these methods, features from source and target are extracted and then the distances between them are calculated using MMD.

3.3 Transfer Learning Model

The transfer learning model is developed based on the basic CNN architecture. Like base model this model also has 2 convolutional layers, 1 max pooling layer and 2 fully connected layers. Each of the convolutional layers has 64 filters and the kernel size is 1x3. The pool size is 1x2 with strides 2 for the max pooling layer. In this transfer learning model both source and target data are passed simultaneously. After extracting the features from both source and target data MMD loss is calculated for each batch of data. The classification loss is obtained from the classifier. The total loss is the summation of MMD loss and classification loss.

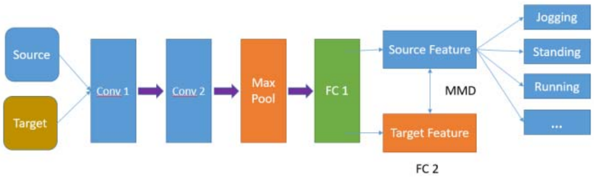

Fig. 2. Transfer learning model

In Fig. 2 the architecture of the transfer learning model is shown that contains convolutional layers, max pooling layer, fully connected layers and the figure also shows that MMD is calculated from source and target features. The transfer learning model calculates the total loss including classification loss and MMD loss and then tries to minimize the loss.

IV. EXPERIEMENT AND EVALUATION

In this section, (i) dataset setup and (ii) evaluation of proposed transfer learning model are presented.

4.1 Dataset Setup

For our experiment, we have used 4 datasets: WISDM, DUMD, HASC and Single Chest Mounted.
DUMD has 10 activities performed by 50 people. Here, for similar activities among datasets we have chosen 4 activities performed by 14 users for this dataset.

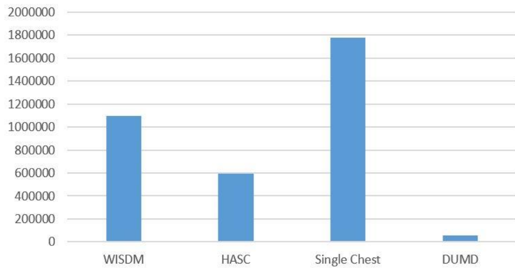

Fig. 3. No. of samples for WISDM, HASC, Single Chest and DUMD datasets

In Fig. 3 different number of samples per dataset is represented. Each dataset contains different number of activities and users. Each dataset contains only accelerometer values of three axes: x, y and z. There are no features like mean, median, max etc. in these datasets. Since CNN is used for classification so it will automatically extract features. Below is a short description of these datasets.

Table 1: Dataset description (A = Accelerometer)

Dataset	Sensors	Users	Activities	Sampling rate
WISDM	A	35	6	20 Hz
Single Chest	A	15	7	52 Hz
HASC	A	10	6	100Hz
DUMD	A	14	4	30 Hz

In table 1, the number of users, sampling rates, activities etc. are shown for different datasets. Each dataset has different number of activities. When two datasets are used for transfer learning, we have used only common activities of two datasets.

4.2 Evaluation of Transfer Learning Model

We have applied transfer learning from one dataset to another and compared the result with base CNN model. First, we have trained base model on one dataset and then tested the model on another dataset, which has similar activities. We have trained the base model on labelled source dataset and then applied it to unlabeled target dataset. For the transfer learning model, we have extracted features from both source and target domain and calculated the MMD loss alongside the classification loss. Here the source data is labeled but the target data is unlabeled. The model tries to minimize the total loss to adapt target dataset. After the completion of training, we have tested the model on target dataset to obtain the test accuracy. Finally, we have compared the performance between base model and transfer learning model based on accuracy. Throughout the experiment, we have used transfer learning model from single source domain to single target domain.

Table 2: Classification accuracy (%) W = WISDM; H = HASC; D = DUMD; S = Single Chest

Dataset	Without Transfer Learning	Using Transfer Learning
W → D	25.34	32.93
W → S	40.67	61.93
H → D	49.96	54.39
S → H	50.46	49.62
S → D	50.68	50.89

In table 2, the accuracies obtained from base model and transfer learning model are shown. Here A→B means we are transferring the knowledge of our model from "A" dataset to "B" dataset.

From the table 2, it is seen that the accuracy gets improved while using transfer learning model than the base CNN model. We have used feature-based transfer learning here. When transferring from one dataset to another dataset, we chose similar type of activities between source and target dataset. Depending on the datasets the number of similar activities were two, three, four etc.

From the experiment, it is found that sometimes for any particular activity of source and target dataset the accuracy degrades. For example, from HASC and DUMD datasets, we chose 3 common activities: "Jogging", "Walking", "Sitting". If we train the base model on HASC dataset for these 3 activities and then test the base model on DUMD dataset, the accuracy becomes 33.47%. For transfer learning model the accuracy is 33.24%. If we remove the "Sitting" activity and then train and test the models then for base model accuracy becomes 49.96% and for transfer learning model it becomes 54.39%.

Similar scenario happens for WISDM and HASC datasets. Therefore, we have visualized and analyzed the signal data of these 3 datasets.

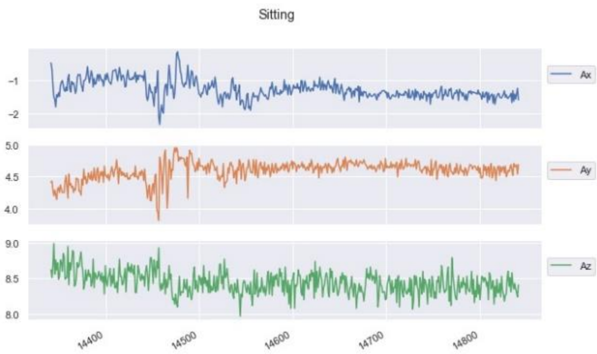

Fig. 4. Sitting activity of DUMD

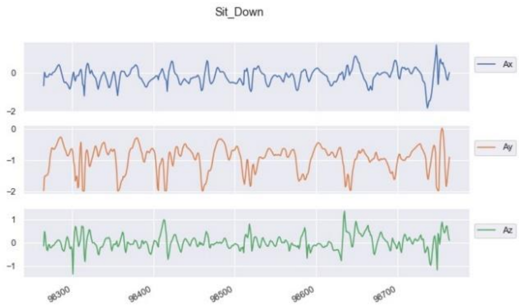

Fig. 5. Sit Down activity of HASC

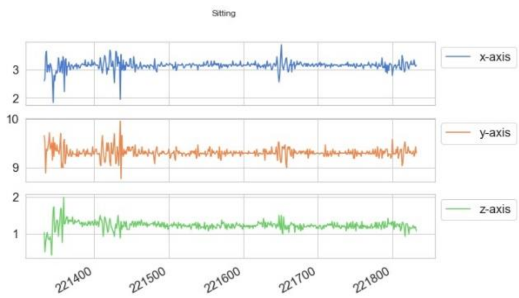

Fig. 6. Sitting activity of WISDM

Fig. 4, 5 and 6 show the graph of "Sitting" activity that contains 3 axes accelerometer data for 3 different datasets.

From the above figures of "Sitting" activity it is seen that the accelerometer value range is different among these datasets. Since the datasets are different, so a particular activity might be performed in different ways. For example, in one dataset the signal value represents the user is sitting idly where in other dataset it probably represents the way of sitting from standing state. This is a big challenge while applying transfer learning from source dataset to target dataset. The problem eventually becomes severe when the target dataset is unlabeled and contains small amount of samples.

When transfer learning is used within the source and target datasets that have similar type of data distribution then the accuracy becomes high. In paper [7] the transfer learning was used from one person to another of same dataset. There each user had same amount of data for each activity. Therefore, when transfer learning was used from one person to another, the model performed well. On the other hand, if number of data for a particular activity is different between 2 datasets then the MMD loss measurement might not proper. If one activity data becomes much larger than the data of similar activity of other dataset then after running several batches the next activity data will come and be compared with the activity of other dataset. In that case, there will be 2 different activities and the MMD loss will be calculated for the feature distribution of these 2 different activities. As a result, the accuracy will be decreased. Therefore, it is also a big challenge to use transfer learning efficiently between source and target domains, if the similar activity of 2 different datasets, contains different number of samples.

V. Discussion

In this paper, we have proposed unsupervised transfer learning based on Maximum Mean Discrepancy (MMD). We have applied the model on 4 datasets, each time making one dataset as source domain and other dataset as target domain. The source domain was labeled but the target domain was unlabeled. We have used base model to train on source dataset and then tested on target dataset. On the other hand, we have passed both labeled source dataset and unlabeled target dataset to transfer learning model. The model extracts the features from both source and target datasets and then uses MMD to calculate discrepancy between the feature distributions. We have compared the performances of both base model and transfer learning model. It is seen that most of the cases the performance of transfer learning model is better than the base model.

In the introduction section, we have listed some key challenges for human activity recognition and tried to solve them using transfer learning. One challenge was insufficient train data. Transfer learning model gathers knowledge from large dataset and transfers knowledge to small dataset. Among the 4 datasets, DUMD has the least number of samples. Therefore, we have trained our transfer learning model on large dataset and then applied the knowledge to DUMD dataset. It is found that the accuracy is improved compared to base model in this scenario.

Another problem was labeled train data collection. Using transfer learning on a labeled train dataset, we can test on unlabeled test dataset, which has similar activities like source dataset. So using transfer learning this problem can be solved.

Finally, probability distribution mismatch. We can use transfer learning model to extract features from source and target datasets. Then we can use MMD to measure the dissimilarity between feature distributions of source and target datasets. We can identify which features are close between source and target datasets. Therefore, using transfer learning model most of the challenges can be solved.

However, the accuracy using transfer learning model is not so high although its performance is better than base model. The reason is large and small number of samples of source and target datasets. If the number of samples for each activity of source and target datasets are not same then features of different activities from source and target datasets might be compared. This will lead to large MMD loss and the accuracy will drop. For better performance the number of samples for each activity from both source and target domains needs to be almost equal. To solve this uneven number of sample problem, one solution might be sorting similar activity samples and then compare similar activity samples. This will prevent comparing different activity samples and the accuracy might improve.

Another cause of relatively low accuracy of transfer learning model is doing same activity of source and target datasets in different ways. The sampling rate, user's age, psychology etc. affect performing same activity in various ways. Therefore, if these type of activities exist in source and target domain then accuracy might drop because due to high variation of feature values the MMD loss will be high. This is a very big challenge for HAR using transfer learning.

VI. Conclusion and Future Work

Collecting large amount of labeled data is time consuming. Transfer learning has been proved effective to solve this problem since it transfers knowledge from large dataset to small dataset, which has little or no labeled data. In this paper, we have proposed unsupervised transfer learning. We use feature-based transfer learning model from one dataset to another different dataset that contain similar type of activities but different number of samples. Although the performance is better than the base model, sometimes it degrades. Different users can perform the same activity in different ways. Therefore, the sensor value pattern for same activity in different dataset varies. As a result, the feature distribution between source and target datasets varies a lot that leads to poor recognition performance for the transfer learning model. In the future work, we will analyze the feature distributions of source and target datasets, which are extracted by deep learning models. This will help us to visualize whether the right features are transferred or not from source to target datasets. We will apply transfer learning among other different datasets and compare performances. Besides, we want to use heterogeneous transfer learning for human activity recognition.

References

[1] T. Hong, J.H.; Ramos, J.; Dey, A.K. Toward Personalized Activity Recognition Systems with a Semipopulation Approach. IEEE Trans. Hum.-Mach. Syst. 2016, 46, 101–112. [CrossRef]

[2] Cao, L.; Liu, Z.; Huang, T.S. Cross-dataset action detection. In Proceedings of the IEEE Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June; pp. 1998–2005.

[3] Z. Zhao, Y.Chen, J. Liu, Z. Shen, M. Liu, Cross-people mobile-phone based actfivity recognition. In Proceedings of the International Joint

Conference on Artificial Intelligence, Barcelona, Spain, 16–22 July 2011; AAAI Press: Palo Alto, CA, USA, 201.

[4] M. Inoue, S. Inoue, T. Nishida, Deep recurrent neural network for mobile human activity recognition with high throughput. arXiv, arXiv:1611.03607, 2016

[5] D. Roggen, Deep convolutional feature transfer across mobile activity recognition domains, sensor modalities and locations. In Proceedings of the ACM International Symposium onWearable Computers, Heidelberg, Germany, 12–16 September; pp. 92–99, 2016

[6] J. Wang, V.W. Zheng, Y. Chen, M. Huang, Deep Transfer Learning for Cross-domain Activity Recognition. In Proceedings of the 3rd International Conference on Crowd Science and Engineering, Singapore, 28–31 July 2018.

[7] D. Renjie, L. Xue, N.Lanshun, N. Jiazhen, S. Xiandong, S. Dianhui, L. Guozhong and Z. Dechen, Empirical Study and Improvement on Deep Transfer Learning for Human Activity Recognition

[8] Khan, Md Abdullah Hafiz; Roy, Nirmaly and Misara, Archan, Scaling human activity recognition via deep learning-based domain adaptation. *.2018 IEEE International Conference on Pervasive Computing and Communications, Athens, Greece, March 19-23*.1-9, 2018.

[9] T. Eric, H. Judy, Z. Ning, S. Kate, and D. Trevor. 2014, Deep domain confusion: Maximizing for domain invariance. arXiv preprint arXiv:1412.3474, 2014.

[10] W. Jindong, C.Yiqiang, H. Lisha, P. Xiaohui, and S. Y. Philip,. Stratifed Transfer Learning for Cross-domain Activity Recognition. In IEEE international conference on pervasive computing and communications (PerCom), 2018

[11] G.Yaroslav, U. Evgeniya, A. Hana, G. Pascal, L. Hugo, L. Franc, M. Mario, and L. Victor. 2016, Domain-adversarial training of neural networks. Journal of Machine Learning Research 17, 59, 1–35, 2016

[12] L. Mingsheng, Z. Han, W. Jianmin, and I. J.Michael, Deep Transfer Learning with Joint Adaptation Networks. In International Conference on Machine Learning. 2208–2217, 2017

[13] G. Arthur, M. Karsten, Borgwardt, J. Malte, J. Rasch, Bernhard Scholkopf, S. Alexander, 2012. A Kernel Two-Sample Test. Journal of Machine Learning Research 13,723-773, 2012

[14] A. Bulling, U. Blanke, and B. Schiele, “A tutorial on human activity recognition using body-worn inertial sensors,” ACM Computing Surveys (CSUR), vol. 46, no. 3, p. 33, 2014.

[15] O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192–1209, 2013.

[16] J.Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, “Deep learning for sensorbased activity recognition: A survey,” arXiv preprint arXiv:1707.03502, 17. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[17] Y. Guan and P.Thomas, Ensembles of deep lstm learners for activity recognition using wearables. IMWUT 1, 2 2017.

[18] H. Shimodaira Improving predictive inference under covariate shift by weighting the log-likelihood function. J Stat Plan Inf. 2000;90(2):227–44, 2000.

[19] R. Gopalan, R. Li, V.M. Patel, R.Chellappa,Domain Adaptation for Visual Recognition. In Foundations and Trends in Computer Graphics; Now Publisher: Delft, Holland, pp. 285–378, 2015

[20] T. Ben, Z. Yu, J. P.Sinno, and Y.Qiang, Distant Domain Transfer Learning. In .irty-First AAAI Conference on Arti€cial Intelligence, 2017.

[21] C.Rita, S.Qian, F. Wei, I.Davidson, S. Panchanathan, and Y. Jieping, Multisource domain adaptation and its application to early detection of fatigue. ACM Transactions on Knowledge Discovery from Data (TKDD) 6, 4 2012.

[22] J. P. Sinno, I. W. Tsang, J. T. Kwok, and Qiang Yang, Domain adaptation via transfer component analysis. IEEE Transactions on Neural Networks 22, 2 2011.